

# VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS

Absender: MIT DER INTERNATIONALEN VORLÄUFIGEN PRÜFUNG BEAUFTRAGTE BEHÖRDE

## PCT

An:

SCHAUMBURG, Karl Heinz  
Postfach 86 07 48  
81634 München  
ALLEMAGNE

**EINGEGANGEN**

02. Okt. 2006

Erled. ....

MITTEILUNG ÜBER DIE ÜBERSENDUNG  
DES INTERNATIONALEN VORLÄUFIGEN  
BERICHTS ZUR PATENTIERBARKEIT

(Regel 71.1 PCT)

Absendedatum  
(Tag/Monat/Jahr)

02.10.2006

Aktenzeichen des Anmelders oder Anwalts  
2003ODT0902 P

### WICHTIGE MITTEILUNG

Internationales Aktenzeichen  
PCT/EP2004/009539

Internationales Anmeldedatum (Tag/Monat/Jahr)  
26.08.2004

Prioritätsdatum (Tag/Monat/Jahr)  
30.09.2003

Anmelder

OCE DOCUMENT TECHNOLOGIES GMBH et al.

1. Dem Anmelder wird mitgeteilt, daß ihm die mit der internationalen vorläufigen Prüfung beauftragte Behörde hiermit den zu der internationalen Anmeldung erstellten internationalen vorläufigen Bericht zur Patentierbarkeit, gegebenenfalls mit den dazugehörigen Anlagen, übermittelt.
2. Eine Kopie des Berichts wird - gegebenenfalls mit den dazugehörigen Anlagen - dem Internationalen Büro zur Weiterleitung an alle ausgewählten Ämter übermittelt.
3. Auf Wunsch eines ausgewählten Amtes wird das Internationale Büro eine Übersetzung des Berichts (jedoch nicht der Anlagen) ins Englische anfertigen und diesem Amt übermitteln.

#### 4. ERINNERUNG

Zum Eintritt in die nationale Phase hat der Anmelder vor jedem ausgewählten Amt innerhalb von 30 Monaten ab dem Prioritätsdatum (oder in manchen Ämtern noch später) bestimmte Handlungen (Einreichung von Übersetzungen und Entrichtung nationaler Gebühren) vorzunehmen (Artikel 39 (1)) (siehe auch die durch das Internationale Büro im Formblatt PCT/IB/301 übermittelte Information).

Ist einem ausgewählten Amt eine Übersetzung der internationalen Anmeldung zu übermitteln, so muß diese Übersetzung auch Übersetzungen aller Anlagen zum internationalen vorläufigen Bericht zur Patentierbarkeit enthalten. Es ist Aufgabe des Anmelders, solche Übersetzungen anzufertigen und den betroffenen ausgewählten Ämtern direkt zuzuleiten.

Weitere Einzelheiten zu den maßgebenden Fristen und Erfordernissen der ausgewählten Ämter sind Band II des PCT-Leitfadens für Anmelder zu entnehmen.

Der Anmelder wird auf Artikel 33(5) hingewiesen, in welchem erklärt wird, daß die Kriterien für Neuheit, erfinderische Tätigkeit und gewerbliche Anwendbarkeit, die im Artikel 33(2) bis (4) beschrieben werden, nur für die internationale vorläufige Prüfung Bedeutung haben, und daß "jeder Vertragsstaat (...) für die Entscheidung über die Patentfähigkeit der beanspruchten Erfindung in diesem Staat zusätzliche oder abweichende Merkmale aufstellen" kann (siehe auch Artikel 27(5)). Solche zusätzlichen Merkmale können z.B. Ausnahmen von der Patentierbarkeit, Erfordernisse für die Offenbarung der Erfindung sowie Klarheit und Stützung der Ansprüche betreffen.

Name und Postanschrift der mit der internationalen Prüfung beauftragten Behörde



Europäisches Patentamt - P.B. 5818 Patentlaan 2  
NL-2280 HV Rijswijk - Pays Bas  
Tel. +31 70 340 - 2040 Tx: 31 651 epo nl  
Fax: +31 70 340 - 3016

Bevollmächtigter Bediensteter

Flaßhar, Claudia

Tel. +31 70 340-4789




# VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS

## PCT

### INTERNATIONALER VORLÄUFIGER BERICHT ÜBER DIE PATENTIERBARKEIT

(Kapitel II des Vertrags über die internationale Zusammenarbeit auf dem Gebiet des Patentwesens)

Aktenzeichen des Anmelders oder Anwalts 2003ODT0902 P	<b>WEITERES VORGEHEN</b> siehe Formblatt PCT/PEA/416	
Internationales Aktenzeichen PCT/EP2004/009539	Internationales Anmeldedatum (Tag/Monat/Jahr) 26.08.2004	Prioritätsdatum (Tag/Monat/Jahr) 30.09.2003
Internationale Patentklassifikation (IPC) oder nationale Klassifikation und IPC INV. G06K9/20		
Anmelder OCE DOCUMENT TECHNOLOGIES GMBH et al.		
<p>1. Bei diesem Bericht handelt es sich um den internationalen vorläufigen Prüfungsbericht, der von der mit der internationalen vorläufigen Prüfung beauftragten Behörde nach Artikel 35 erstellt wurde und dem Anmelder gemäß Artikel 36 übermittelt wird.</p> <p>2. Dieser BERICHT umfaßt insgesamt 11 Blätter einschließlich dieses Deckblatts.</p> <p>3. Außerdem liegen dem Bericht ANLAGEN bei; diese umfassen</p> <p>a. <input checked="" type="checkbox"/> (an den Anmelder und das Internationale Büro gesandt) insgesamt 6 Blätter; dabei handelt es sich um</p> <p><input checked="" type="checkbox"/> Blätter mit der Beschreibung, Ansprüchen und/oder Zeichnungen, die geändert wurden und diesem Bericht zugrunde liegen, und/oder Blätter mit Berichtigungen, denen die Behörde zugestimmt hat (siehe Regel 70.16 und Abschnitt 607 der Verwaltungsvorschriften).</p> <p><input type="checkbox"/> Blätter, die frühere Blätter ersetzen, die aber aus den in Feld Nr. 1, Punkt 4 und im Zusatzfeld angegebenen Gründen nach Auffassung der Behörde eine Änderung enthalten, die über den Offenbarungsgehalt der internationalen Anmeldung in der ursprünglich eingereichten Fassung hinausgeht.</p> <p>b. <input type="checkbox"/> (nur an das Internationale Büro gesandt) insgesamt (bitte Art und Anzahl der/des elektronischen Datenträger(s) angeben), der/die ein Sequenzprotokoll und/oder die dazugehörigen Tabellen enthält/enthalten, nur in elektronischer Form, wie im Zusatzfeld betreffend das Sequenzprotokoll angegeben (siehe Abschnitt 802 der Verwaltungsvorschriften).</p>		
<p>4. Dieser Bericht enthält Angaben zu folgenden Punkten:</p> <p><input checked="" type="checkbox"/> Feld Nr. I Grundlage des Berichts</p> <p><input type="checkbox"/> Feld Nr. II Priorität</p> <p><input type="checkbox"/> Feld Nr. III Keine Erstellung eines Gutachtens über Neuheit, erfinderische Tätigkeit und gewerbliche Anwendbarkeit</p> <p><input type="checkbox"/> Feld Nr. IV Mangelnde Einheitlichkeit der Erfindung</p> <p><input checked="" type="checkbox"/> Feld Nr. V Begründete Feststellung nach Artikel 35(2) hinsichtlich der Neuheit, der erfinderischen Tätigkeit und der gewerblichen Anwendbarkeit; Unterlagen und Erklärungen zur Stützung dieser Feststellung</p> <p><input type="checkbox"/> Feld Nr. VI Bestimmte angeführte Unterlagen</p> <p><input type="checkbox"/> Feld Nr. VII Bestimmte Mängel der internationalen Anmeldung</p> <p><input type="checkbox"/> Feld Nr. VIII Bestimmte Bemerkungen zur internationalen Anmeldung</p>		
Datum der Einreichung des Antrags  13.04.2005	Datum der Fertigstellung dieses Berichts  02.10.2006	
Name und Postanschrift der mit der internationalen vorläufigen Prüfung beauftragten Behörde   Europäisches Patentamt - P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk - Pays Bas Tel. +31 70 340 - 2040 Tx: 31 651 epo nl Fax: +31 70 340 - 3016	Bevollmächtigter Bediensteter  Neubüser, Bernhard  Tel. +31 70 340-4983	



---

**Feld Nr. I Grundlage des Berichts**

---

1. Hinsichtlich der **Sprache** beruht der Bescheid auf

- ☒ der internationalen Anmeldung in der Sprache, in der sie eingereicht wurde.
- ☐ einer Übersetzung der internationalen Anmeldung in die folgende Sprache, bei der es sich um die Sprache der Übersetzung handelt, die für folgenden Zweck eingereicht worden ist:
  - ☐ internationale Recherche (nach Regeln 12.3 a) und 23.1 b))
  - ☐ Veröffentlichung der internationalen Anmeldung (nach Regel 12.4 a))
  - ☐ internationale vorläufige Prüfung (nach Regeln 55.2 a) und/oder 55.3 a))

2. Hinsichtlich der **Bestandteile\*** der internationalen Anmeldung beruht der Bericht auf *(Ersatzblätter, die dem Anmeldeamt auf eine Aufforderung nach Artikel 14 hin vorgelegt wurden, gelten im Rahmen dieses Berichts als "ursprünglich eingereicht" und sind ihm nicht beigelegt)*:

**Beschreibung, Seiten**

1-18 in der ursprünglich eingereichten Fassung  
3a, 3b eingegangen am 11.07.2005 mit Schreiben vom 11.07.2005

**Ansprüche, Nr.**

1-19 eingegangen am 11.07.2005 mit Schreiben vom 11.07.2005

**Zeichnungen, Blätter**

1/8-8/8 in der ursprünglich eingereichten Fassung

- ☐ einem Sequenzprotokoll und/oder etwaigen dazugehörigen Tabellen - siehe Zusatzfeld betreffend das Sequenzprotokoll

3. ☐ Aufgrund der Änderungen sind folgende Unterlagen fortgefallen:

- ☐ Beschreibung: Seite
- ☐ Ansprüche: Nr.
- ☐ Zeichnungen: Blatt/Abb.
- ☐ Sequenzprotokoll (*genaue Angaben*):
- ☐ etwaige zum Sequenzprotokoll gehörende Tabellen (*genaue Angaben*):

4. ☐ Dieser Bericht ist ohne Berücksichtigung (von einigen) der diesem Bericht beigelegten und nachstehend aufgelisteten Änderungen erstellt worden, da diese aus den im Zusatzfeld angegebenen Gründen nach Auffassung der Behörde über den Offenbarungsgehalt in der ursprünglich eingereichten Fassung hinausgehen (Regel 70.2 c)).

- ☐ Beschreibung: Seite
- ☐ Ansprüche: Nr.
- ☐ Zeichnungen: Blatt/Abb.
- ☐ Sequenzprotokoll (*genaue Angaben*):
- ☐ etwaige zum Sequenzprotokoll gehörende Tabellen (*genaue Angaben*):

\* Wenn Punkt 4 zutrifft, können einige oder alle dieser Blätter mit der Bemerkung "ersetzt" versehen werden.

---

**Feld Nr. V Begründete Feststellung nach Artikel 35 (2) hinsichtlich der Neuheit, der erfinderischen Tätigkeit und der gewerblichen Anwendbarkeit; Unterlagen und Erklärungen zur Stützung dieser Feststellung**

---

1. Feststellung

Neuheit (N) Ja: Ansprüche 1-19

Nein: Ansprüche

Erfinderische Tätigkeit (IS) Ja: Ansprüche

Nein: Ansprüche 1-19

Gewerbliche Anwendbarkeit (IA) Ja: Ansprüche: 1-19

Nein: Ansprüche:

2. Unterlagen und Erklärungen (Regel 70.7):

**siehe Beiblatt**

**Zu Punkt V.**

- 1 Im vorliegenden Bescheid wird auf folgende Dokumente verwiesen:  
D<sub>1</sub>: US 5 317 646 A (SANG JR HENRY W ET AL) 31. Mai 1994 (1994-05-31)  
D<sub>2</sub>: US 2002/141660 A1 (PUCCI JORGE PABLO ET AL) 3. Oktober 2002  
(2002-10-03)  
D<sub>3</sub>: US 6 028 970 A (DIPIAZZA PHILIP SILVANO ET AL) 22. Februar 2000  
(2000-02-22)
- 2 Der Gegenstand des Anspruchs 1 beruht nicht auf einer erfinderischen Tätigkeit im Sinne von Artikel 33(3) PCT.
- 2.1 Dokument D<sub>2</sub> wird als nächstliegender Stand der Technik angesehen und offenbart (die Verweise in Klammern beziehen sich auf dieses Dokument; Durchstreichungen (~~wie hier~~) markieren Passagen aus Anspruch 1, die keine Entsprechung in D<sub>2</sub> haben):

Ein Verfahren zum Erfassen von Daten aus maschinell lesbaren Dokumenten, ~~wobei die Daten einer Datenbank zugeordnet werden~~, indem einzelne Daten möglichst automatisch dem Dokument extrahiert werden und in entsprechende ~~Datenbankfelder~~ eingetragen werden,

(Zusammenfassung: "The document scanner, system and method operates in conjunction with a document imprinted with data and a plurality of form documents adapted to have data imprinted thereon. The documents have at least one and typically many data image fields. Ultimately, the document scanner, system and method output a delimited string of decoded characters to another computer system via a common computer communications port. ... The system selects one of the stored forms, extracts the data from each data field, decodes or calculated the data, and validates the data (in the presenece of data validation parameters) and stores the decoded/calculated data.", par. [0052], Ende: "It should be appreciated that the further computer device can easily process this delimied string of decoded characters into a spreadsheet, database or any other type of word processing program.")

und falls Daten für ein oder mehrere bestimmte ~~Datenbankfelder~~ einem Dokument nicht mit der notwendigen Zuverlässigkeit extrahiert werden konnten  
(Zusammenfassung, Ende: "A data reporting and data correction system, activated in the

presence of the data error reporting and correction descriptor, enables correction of errors"),  
ausführen folgender Schritte:

- Darstellen des Dokuments an einem Bildschirm (implizit),
- Anzeigen des Datenbankfeldes, für das die Daten nicht mit der notwendigen Zuverlässigkeit extrahiert werden konnten, am Bildschirm,

(Par. [0051]: "Any error reports from field and rule checker unit 62 are supplied via control unit 80 to display 84. The operator at keyboard 86 may correct the error if the data correction field descriptor has been turn[ed] ON. If the operator is enabled to correct the data and does correct the error, summation module 64 substitutes the corrected data for the previously scanned and decoded incorrect data.")

~~- Ausführen einer Vorschlags-Routine, mit welcher Stringabschnitte in der Nähe eines von einem Benutzer auf dem Bildschirm bewegbaren Zeiger ausgewählt, markiert und zur Extraktion vorgeschlagen werden.~~

2.2 Der Gegenstand von Anspruch 1 unterscheidet sich von der Lehre von D<sub>2</sub> in folgenden Punkten:

- i) Die zu erfassenden Daten sind einer Datenbank zugeordnet.
- ii) Die erfassten Daten werden in Datenbankfelder eingetragen.
- iii) Im Falle, dass (Datenbank-)Felder nicht mit der notwendigen Zuverlässigkeit extrahiert werden konnten, Ausführen einer Vorschlags-Routine, mit welcher Stringabschnitte in der Nähe eines von einem Benutzer auf dem Bildschirm bewegbaren Zeiger ausgewählt, markiert und zur Extraktion vorgeschlagen werden.

2.3 Die genannten Unterschiede lassen viele Interpretationen zu. Die Unterschiede wurden wie folgt interpretiert (gleiche Ziffern i)-iii) korrespondieren zueinander):

- i) Es ist eine Möglichkeit der Speicherung der Daten in einer Datenbank vorgesehen, i.e. dem Verfahren sind eine Datenbank und eine Liste von Feldern der Datenbank mitsamt Korrespondenzen zu Feldern, die aus dem Dokument extrahiert werden, bekannt.
- ii) Die erfassten Daten werden in Komponenten einer Datenstruktur gespeichert, die notwendigerweise verwendet werden muss, wenn die Daten in einer Datenbank gespeichert werden soll, d.h. ii) wird durch i) impliziert.

- iii) Als eine "Vorschlags-Routine, mit welcher Stringabschnitte in der Nähe eines von einem Benutzer auf dem Bildschirm bewegbaren Zeiger ausgewählt, markiert und zur Extraktion vorgeschlagen wird" wird auch eine Routine zur Steuerung eines "Maus"-Zeigers angesehen, die die Auswahl von Stringabschnitten, z.B. durch Definition eines rechteckigen Bildschirmbereichs, gestattet.

Es wird darauf hingewiesen, dass nach Auffassung des Prüfers aus der gewählten Formulierung nicht abgeleitet werden kann, dass die Vorschlagsroutine mit erkannten (alphanumerischen) Daten arbeitet (und nur deren Position in der Bitmap benutzt) und nicht, wie in D<sub>2</sub>, durch ein pixelweises Vorgehen Bildbereiche definiert und mittels OCR Stringabschnitte extrahiert.

Es wird darauf hingewiesen, dass diese Auffassung *zusätzlich* durch die Formulierung von Anspruch 8 gestützt wird, gemäß der "und die Vorschlagsroutine *zusätzlich* zum markierten Stringabschnitt in grafischer Darstellung den codierten Text dieses Stringabschnittes darstell."

- 2.4 Die mit der vorliegenden Erfindung zu lösende Aufgabe kann somit darin gesehen werden, daß

- i), ii) Eine Möglichkeit der Speicherung der extrahierten Daten in einer Datenbank (i.e. einer strukturierten, dauerhaften, durchsuchbaren Datenablage) zu schaffen.

- iii) Ein bequeme Eingabemöglichkeit für Korrekturen im Falle von Fehlern oder unsicheren Ergebnissen bei der Datenextraktion zu schaffen.

- 2.5 Die Unterschiede/Aufgaben i), ii) einerseits und iii) andererseits sind vollkommen unabhängig voneinander, die zugehörigen Merkmale stellen also eine Aneinanderreihung dar.

- 2.6 Bei der Beurteilung des Vorliegens eines erfinderischen Schritts sind daher i), ii) einerseits, und iii) andererseits unabhängig und getrennt voneinander zu betrachten.

- 2.7 Zu i), ii): Die Möglichkeit der Speicherung der Daten in einer Datenbank wird schon in D<sub>1</sub> in Erwägung gezogen (Par. [0051], Ende: "It should be appreciated that the further

computer device can easily process this delimited string of decoded characters into a spreadsheet, database ..."). Den Abschnitten [0036]-[0041] ist zu entnehmen, dass bei dem Verfahren gemäß  $D_2$  die zur Datenbankankbindung nötigen Feldinformation (Metadaten) vorgehalten werden. Da Anspruch 1 keine weiteren Merkmale bzgl. der Datenbankankbindung enthält, wird die genannte Passage von  $D_2$  als ausreichender Hinweis für den Fachmann angesehen, die Aspekte i), ii) von Anspruch 1 aus  $D_2$  abzuleiten.

Davon abgesehen, zeigen Dokument  $D_1$  (Fig. 2, 700: "Database Insertion") oder  $D_3$  (Fig. 1B), dass das Abspeichern von aus Dokumenten extrahierten Daten in einer Datenbank bekannt ist.

Zu iii): Wie in  $D_2$  in den Paragraphen [0050] und [0051] dargestellt ist, werden die extrahierten Daten geprüft. Im Falle eines gefundenen Fehlers besteht die Möglichkeit der manuellen Korrektur. Wie die manuelle Korrektur genau von Statten geht, darüber schweigt  $D_2$ .

Wenn ein Fachmann eine für den Benutzer einfache und bequeme Methode der manuellen Korrektur von Fehlern zu implementieren hat, wird er erkennen, dass notwendigerweise angezeigt werden muss,

- für welches Feld ungültige Daten extrahiert worden sind,
- woher die Daten stammen (in der Bitmap des eingescannten Dokuments).

Es wird anerkannt, dass der Fachmann in Betracht ziehen könnte, nur die fehlerhaften Daten und den Bereich der Bitmap, aus dem fehlerhafte Daten extrahiert worden sind, anzuzeigen. Offensichtlich müßte der Nutzer des Verfahrens dann die Daten manuell über die Tastatur eingeben. Es wird aber erachtet, dass der Fachmann sicher auch in Betracht ziehen würde eine einfache Möglichkeit zu schaffen, mit der ein Bereich/Bereiche der Bitmap des Dokuments ausgewählt wird/werden, und die ohnehin vorhandene OCR-Funktionalität zur Extraktion der Daten aus der Bitmap verwendet wird.

Damit stimmt die benötigte Funktionalität mit der für die Felddefinition benötigten Funktionalität überein (und kann also - zumindest teilweise - wiederverwendet werden). Gemäß Abschnitt [0042] von  $D_2$  wird ein Zeiger verwendet, um Position und Größe der Felder zu definieren.



Daher wird erachtet, dass dem Fachmann alleine schon aus der Lehre von D<sub>2</sub> auch die Merkmale iii) nahegelegt werden.

Davon abgesehen ist dem Fachmann mit D<sub>1</sub> ein Verfahren bekannt, dass eine besonders einfache Definition von Teilen eines Dokuments als zu extrahierende Felder beschreibt. Im Gegensatz zu dem durch D<sub>1</sub> nahegelegten Verfahren braucht bei diesem Verfahren nicht einmal die Größe des zu extrahierenden Bereichs manuell festgelegt zu werden.

- 3 Der unabhängige Anspruch 12<sup>1</sup> genügt nicht den Erfordernissen von Artikel 6 PCT. Dem Gegenstand von Anspruch 12 scheint kein erfinderischer Schritt im Sinne von Artikel 33(3) PCT zugrunde zu liegen.
- 3.1 Die Passage Zeilen 20-27 von Anspruch 12 wird so verstanden, dass durch einen Vergleich *des Inhalts* eines Stringabschnitts der weiter unten in einer Tabelle vorkommt mit Stringabschnitten, die in der ersten Zeile vorkommen, ermittelt wird, in welches Feld der Stringabschnitt extrahiert werden muss.

Ein solches Vorgehen liefert jedoch nur in Ausnahmefällen das gewünschte Resultat. Es müssen dazu nämlich in den Spalten fast übereinstimmende Einträge stehen und die verschiedenen Spalten müssen sich deutlich voneinander unterscheiden. Der Leser wird daher Zweifel haben, welches der angestrebte Schutzbereich ist.

Die angegebene Interpretation ist ferner nicht im Einklang mit der Beschreibung. Die Extraktion von Daten aus Tabellen ist auf S. 12, Zeile 1 - S. 15, Zeile 10 beschrieben; die Passage S. 12, Zeile 31 - S. 13, Zeile 7 beschreibt den Vergleich von Stringabschnitten mit Hilfe einer Kostenfunktion.

Gemäß der letztgenannten Passage findet ein Vergleich der *horizontalen Position und der Breite* der Stringabschnitte statt.

Die gewählte Formulierung von Anspruch 12 wird daher als irreführend und nicht

---

<sup>1</sup> Wegen der Verwendung von "insbesondere" im Nebensatz "insbesondere nach einem der Ansprüche 1 - 11" kann Anspruch 12 nicht als *notwendigerweise* von einem der Ansprüche 1 - 11 abhängig betrachtet werden.

durch die Beschreibung gestützt angesehen.

- 3.2 Dokument D<sub>3</sub> offenbart, dass Regeln erstellt werden, die Datenlisten oder Tabellen beschreiben und zur Extraktion und Fehleranalyse/Korrektur verwendet werden (Spalte 2, Zeilen 38-45: "A third type of rule is a position verifier. This type of rule requires that certain ordering logic inherent in the definition of the data fields by followed ...", Spalte 13, Zeilen 36-40, Fig. 4, Fig. 5). Eine manuelle Korrektur ist vorgesehen, s. Fig. 1B, 1C: "Operator review, if required".

Daher wird der Gegenstand von Anspruch 12, wie er im Licht der Beschreibung verstanden wird, als durch D<sub>3</sub> nahegelegt angesehen.

- 3.3 Es wird darauf hingewiesen, dass der Gegenstand von Anspruch 12 selbst dann als nicht erfinderisch angesehen werden kann, wenn das Wort "insbesondere" in Anspruch 12 gestrichen wird (und also Anspruch 12 als vom Anspruch 1 abhängig gelesen wird), da gemäß obigem Absatz 2 das Dokument D<sub>2</sub> den Gegenstand von Anspruch 1 nahelegt und die nicht aus D<sub>2</sub> bekannten Merkmale von Anspruch 12 unabhängig (im Sinne einer Aneinanderreihung) sind von den Merkmalen von Anspruch 1, die aus D<sub>2</sub> bekannt sind. Die nicht aus D<sub>2</sub> bekannten Merkmale von Anspruch 12 werden aber von D<sub>3</sub> nahegelegt.
- 4 Der Gegenstand des Anspruchs 16 beruht nicht auf einer erfinderischen Tätigkeit im Sinne von Artikel 33(3) PCT, weil der Gegenstand der Ansprüche 1 und 12 auf keiner erfinderischen Tätigkeit beruht und die gemäß Anspruch 16 verwendeten Hardwarekomponenten allgemein üblich sind.
- 5 Der Gegenstand des Anspruchs 19 beruht nicht auf einer erfinderischen Tätigkeit im Sinne von Artikel 33(3) PCT, weil der Gegenstand der Ansprüche 1 und 12 auf keiner erfinderischen Tätigkeit beruht und davon auszugehen ist, dass das Verfahren aus Dokument D<sub>2</sub> als Computerprogrammprodukt implementiert ist.

**6 ABHÄNGIGE ANSPRÜCHE 2-11, 13-15, 17, 18**

Die Ansprüche 2-11, 13-15, 17, 18 enthalten keine Merkmale, die in Kombination mit den Merkmalen irgendeines Anspruchs, auf den sie sich beziehen, die Erfordernisse des PCT in Bezug auf Neuheit bzw. erfinderische Tätigkeit erfüllen.

6.1 Der Gegenstand der Ansprüche 2, 3, 4 und 5 wird durch D<sub>2</sub> (Fig. 1: "Dictionaries (64)", "Logical Check (68)", "Fields & Interfield Rule Checker (62)") nahegelegt. Ferner wird darauf hingewiesen, dass der Term "Konzept-Informationen" üblicherweise nicht als Oberbegriff für syntaktische und semantische Informationen verwendet wird. Anspruch 4 ist daher nicht im Sinne von Artikel 6 PCT klar.

6.2 Die zusätzlichen Merkmale der Ansprüche 6, 7, 9-11 sind trivial.

6.3 Zu Anspruch 8: Dokument D<sub>2</sub> lehrt, dass entdeckte Fehler manuell korrigiert werden können, s. Par. [0051]: "Any error reports from field and rule checker unit 62 are supplied via control unit 80 to display 84. The operator at keyboard 86 may correct the error if the data correction field descriptor has been turned ON.". Damit der Benutzer die Korrektur vornehmen kann, muss der erkannte Text auch dargestellt werden.

Daher werden die zusätzlichen Merkmale von Anspruch 8 als durch D<sub>2</sub> nahegelegt angesehen.

**6.4 Zu Anspruch 13**

i. Der Ausdruck "String Matching-Verfahren" beschreibt eine große Klasse von Verfahren. Häufig werden hiermit Verfahren zur Abstimmung von (alpha-numerischen) Zeichenketten bezeichnet. Obwohl in dem in Anspruch 13 beschriebenen Verfahren alpha-numerische Zeichenketten verarbeitet werden, bezieht sich das Auffinden von (partiellen) Übereinstimmungen (engl. matching) nicht auf alpha-numerische Zeichen, sondern auf rechteckige Bildausschnitte, deren Ähnlichkeit anhand des Grades der Übereinstimmung von der Position bzw. der Größe bestimmt wird (und nicht anhand der in diesen Bereichen vorgefundenen Zeichenfolgen). Daher ist der Ausdruck "String Matching-Verfahren" irreführend und Anspruch 13 ist nicht klar (Artikel 6 PCT).

Sehr wohl in Übereinstimmung mit der üblichen Bedeutung von "String Matching-Verfahren" ist die Passage S. 14, Zeilen 4-9 der Beschreibung der vorliegenden Anmeldung. Diese bezieht sich aber auf die Prüfung der Konsistenz eines Stringabschnitts mit Konzeptinformationen und nicht auf den Vergleich von zwei Stringabschnitten, die aus dem Dokument extrahiert worden. Sollte beabsichtigt gewesen sein, dass das "String Matching-Verfahren" sich auf diese Passage der Beschreibung bezieht, so hätte eine deutlichere Formulierung gewählt werden müssen.

- ii. Der Anspruch 13 spezifiziert nicht in welcher Weise das String Matching-Verfahren bei dem definierten Verfahren verwendet wird.
  - iii. Dokument D<sub>3</sub> offenbart, dass Regeln (dort "rules") für die Position von Feldern verwendet werden und das Regeln zusammengesetzt werden können. Dies wird als eine Art von String Matching-Verfahren (in dem Sinne, wie es in der Beschreibung verwendet wird: Vergleich von Position und/oder Breite von Bildbereichen, die die Stringabschnitte darstellen). Die zusätzlichen Merkmale von Anspruch 13 sind daher aus D<sub>3</sub> bekannt.
- 6.5 Das zusätzliche Merkmal von Anspruch 14 ist trivial.
- 6.6 Wie in obigem Abschnitt 3.2 dargestellt, offenbart D<sub>3</sub> eine manuelle Korrekturmöglichkeit. Diese ist äquivalent zu den in Anspruch 15 spezifizierten "Funktionen zum Editieren", weshalb der Gegenstand von Anspruch 15 durch D<sub>3</sub> nahegelegt wird.
- 6.7 Die zusätzlichen Merkmale der Ansprüche 17 und 18 sind trivial und ausserdem aus D<sub>2</sub> bekannt.

- 3a -

Aus der US 5,317,646 geht ein Verfahren hervor, das einen Bediener beim Erzeugen von elektronischen Vorlagen für ein Formularerkennungssystem unterstützt. Hierzu wird ein nicht mit Daten versehenes Formular, ein sogenanntes Master-Formular, an einem Bildschirm dargestellt und der Benutzer kann die Datenfelder mit einer Zeigereinrichtung kennzeichnen. Die Koordinaten, die den entsprechenden Bereich begrenzen, werden automatisch erfasst, nachdem ein einziger Punkt innerhalb dieses Bereiches vom Bediener ausgewählt worden ist. Mit diesem Verfahren können schnell und einfach Vorlagen für die automatische Formularerkennung erstellt werden.

In Casey R.G. et al., „Intelligent Forms Processing“, IBM Systems Journal vol. 29 (1990) Nr. 3, Seiten 435 bis 450 wird ein Formularerkennungsverfahren beschrieben, bei dem ein eingescanntes Formular mittels Bildverarbeitungstechniken analysiert und mit anderen gespeicherten Vorlageformularen verglichen wird. Falls keine Übereinstimmung mit einem Vorlageformular gefunden wird, muss ein neues Vorlageformular durch Eingabe am Computer erzeugt werden. Beim Erzeugen einer Vorlage wird das gescannte Formular am Bildschirm dargestellt und mit einer Zeigereinrichtung werden die Begrenzungslinien der Eingabefelder markiert.

Aus der US 2002/141660 A1 geht ein zweistufiges Verfahren hervor, bei welchem zunächst Formularvorlagen eingegeben werden können und danach anhand der eingegebenen Formularvorlagen Dokumente automatisch ausgelesen werden können. Einzugebende Formularvorlagen werden gescannt und der Bediener zeigt mit einem Cursor auf Eingabefelder. Die Position und Größe der Eingabefelder wird gespeichert. Der Bediener kann auch den jedem Datenfeld zugeordneten Datentyp bestimmen. Beim automatischen Lesen von Formularen werden diese eingescannt und

- 3b -

anhand der in den gespeicherten Formularvorlagen enthaltenen Datenfeldern automatisch ausgelesen. Falls beim Auslesen ein Fehler auftritt, kann der Bediener über die Tastatur den Fehler korrigieren.

5

Die US 6,028,970 betrifft ein Verfahren und ein System zur automatischen Texterkennung (OCR). Das System umfasst ein Fehlerkorrigiermodul („error correction logic module“). Dieses Fehlerkorrigiermodul wird auf klar erkennbare Datenfehler angewandt, um diese zu korrigieren. Diese Korrekturen werden automatisch ausgeführt. Hierbei werden nicht nur Fehler einzelner Buchstaben erkannt, sondern auch Fehler im Kontext analysiert und entsprechend korrigiert. Ein nicht automatisch korrigierbarer Fehler kann dem Bediener mittels einer Fehler-

10

15

nachricht mitgeteilt werden. Der Bediener kann dann den mittels der Texterkennung erzeugten Text beurteilen und gegebenenfalls korrigieren.

20

## Ansprüche

1. Verfahren zum Erfassen von Daten aus maschinell lesbaren Dokumenten, wobei die Daten einer Datenbank zugeordnet werden, indem einzelne Daten möglichst automatisch dem Dokument extrahiert werden und in entsprechende Datenbankfelder eingetragen werden, und falls Daten für ein oder mehrere bestimmte Datenbankfelder einem Dokument nicht mit der notwendigen Zuverlässigkeit extrahiert werden konnten, werden folgende Schritte ausgeführt:
- Darstellen des Dokumentes an einem Bildschirm,
  - Anzeigen des Datenbankfeldes, für das die Daten nicht mit der notwendigen Zuverlässigkeit extrahiert werden konnten, am Bildschirm,
  - Ausführen einer Vorschlags-Routine, mit welcher Stringabschnitte in der Nähe eines von einem Benutzer auf dem Bildschirm bewegbaren Zeiger ausgewählt, markiert und zur Extraktion vorgeschlagen werden.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass der Stringabschnitt nach Maßgabe von dem Datenbankfeld zugeordneten Konzept-Informationen ausgewählt, markiert und zur Extraktion vorgeschlagen wird.
3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, dass die Konzept-Informationen die Syntax und/oder die Semantik des Datenbankfeldes beschreiben, so dass von der Vorschlags-Routine ein zu markierender Stringabschnitt entsprechend der Syntax bzw. der Semantik des jeweiligen Datenbankfeldes ausgewählt und markiert wird.

4. Verfahren nach Anspruch 3,  
dadurch gekennzeichnet,  
dass die Informationen zur Syntax die Anzahl von Ziffern  
und/oder Buchstaben und/oder vorbestimmte Formate des zu  
lesenden Stringabschnitts beschreiben.
5. Verfahren nach Anspruch 3 oder 4,  
dadurch gekennzeichnet,  
dass die Informationen zur Semantik vorbestimmte Begriffe,  
bspw. mit einem Lexikon, beschreiben.
6. Verfahren nach einem der Ansprüche 1 bis 5,  
dadurch gekennzeichnet,  
dass ein Stringabschnitt ausgewählt wird, der zwischen  
zwei Begrenzungszeichen angeordnet ist.
7. Verfahren nach Anspruch 6,,  
dadurch gekennzeichnet,  
dass Begrenzungszeichen Leerzeichen und/oder Satzzeichen  
umfassen.
8. Verfahren nach einem der Ansprüche 1 bis 7,  
dadurch gekennzeichnet,  
dass der Text von Dokumenten in grafischer Darstellung  
zunächst mit einem OCR-Verfahren in codierten Text umgesetzt  
wird und die Vorschlags-Routine zusätzlich zum markierten  
Stringabschnitt in grafischer Darstellung den codierten Text  
dieses Stringabschnittes darstellt.
9. Verfahren nach einem der Ansprüche 1 bis 7,  
dadurch gekennzeichnet,  
dass zusätzlich zum markierten Stringabschnitt dieser  
Stringabschnitt nochmals in vergrößerter Darstellung am  
Bildschirm angezeigt wird.



10. Verfahren nach einem der Ansprüche 1 bis 9,  
dadurch gekennzeichnet,  
dass die Vorschlags-Routine nach dem Markieren eines  
Stringabschnittes eine Funktion aktiviert, mit welcher  
durch Betätigen einer oder mehrerer vorbestimmter Tasten  
der Inhalt des markierten Stringabschnittes in die Daten-  
bank übernommen wird.
11. Verfahren nach einem der Ansprüche 1 bis 10,  
dadurch gekennzeichnet,  
dass beim Ausführen der Vorschlags-Routine nach dem Bewe-  
gen des Zeigers ein vorbestimmtes Zeitintervall abgewart-  
et wird, innerhalb dessen der Zeiger nicht bewegt werden  
darf, bevor ein Stringabschnitt ausgewählt wird.
12. Verfahren zum Erfassen von Daten aus maschinell lesbaren  
Dokumenten, wobei die Daten einer Datenbank zugeordnet  
werden, insbesondere nach einem der Ansprüche 1 - 11,  
dadurch gekennzeichnet,  
dass nach dem Einlesen von Daten aus einer ersten Tabel-  
lenzeile in korrespondierende Datenbankfelder automatisch  
die weiteren Tabelleneinträge durch einen Vergleich von  
unterhalb der ersten Tabellenzeile angeordneten Stringab-  
schnitten mit den Stringabschnitten der ersten Tabellen-  
zeile, die den eingelesenen Daten entsprechen, ermittelt  
werden und diese weiteren Tabelleneinträge automatisch  
extrahiert werden.
13. Verfahren nach Anspruch 12,  
dadurch gekennzeichnet,  
dass der Vergleich zwischen den Stringabschnitten mittels  
eines String Matching-Verfahrens erfolgt.
14. Verfahren nach Anspruch 12 oder 13,  
dadurch gekennzeichnet,  
dass die ermittelten Tabelleneinträge markiert werden.

15. Verfahren nach Anspruch 14,  
d a d u r c h g e k e n n z e i c h n e t,  
dass Funktionen zum Editieren der markierten Tabellenein-  
träge bereitgestellt werden.
- 5
16. System zum Erfassen von Daten aus maschinell lesbaren Do-  
kumenten umfassend einen Computer (12) mit einer Spei-  
chereinrichtung (13) und eine CPU (14), wobei in der  
Speichereinrichtung (13) ein Softwareprodukt zum Ausfüh-  
ren des Verfahrens nach einem der Ansprüche 1 - 15 ge-  
speichert ist.
- 10
17. System nach Anspruch 16,  
d a d u r c h g e k e n n z e i c h n e t,  
dass das System eine Eingabeeinrichtung (17) in Form ei-  
ner Maus (6) und/oder Tastatur (7) aufweist.
- 15
18. System nach Anspruch 16 oder 17,  
d a d u r c h g e k e n n z e i c h n e t,  
dass das System einen Scanner (16) zum optischen Abtasten  
von Dokumenten aufweist.
- 20
19. Computerprogrammprodukt, das bei seinem Laden und Ausfüh-  
ren auf einem Computer (12) ein Verfahren nach einem der  
Ansprüche 1 bis 15 bewirkt.
- 25